Weighted Model Counting Without Parameter Variables

No Author Given

No Institute Given

Abstract. The abstract should briefly summarize the contents of the paper in 150–250 words.

Keywords: First keyword · Second keyword · Another keyword.

1 Introduction

Notation. For any propositional formula ϕ over a set of variables X and $p, q \in \mathbb{R}$, let $[\phi]^p_q \colon 2^X \to \mathbb{R}$ be a pseudo-Boolean function defined as

$$[\phi]^p_q(Y) \coloneqq \begin{cases} p & \text{if } Y \models \phi \\ q & \text{otherwise} \end{cases}$$

for any $Y \subseteq X$.

Definition 1. *Let $f, g \colon 2^X \to \mathbb{R}$ be pseudo-Boolean functions. Operations such as addition and multiplication are defined pointwise as*

$$(f+g)(Y) \coloneqq f(Y) + g(Y), \quad \textit{and} \quad (fg)(Y) \coloneqq f(Y)g(Y)$$

for all $Y \subseteq X$. This means that binary operations on pseudo-Boolean functions inherit properties such as associativity and commutativity.

Definition 2. *Let $f \colon 2^X \to \mathbb{R}$ be a pseudo-Boolean function, and $x \in X$. Then $f|_{x=0}, f|_{x=1} \colon 2^X \to \mathbb{R}$ are* restrictions *of f defined as*

$$f|_{x=0}(Y) \coloneqq f(Y \setminus \{x\}), \quad \textit{and} \quad f|_{x=1}(Y) \coloneqq f(Y \cup \{x\})$$

for all $Y \subseteq X$.

Definition 3. *Let X be a set. For any $x \in X$,* projection *$\exists_x$ is an endomorphism $\exists_x \colon \mathbb{R}^{2^X} \to \mathbb{R}^{2^X}$ defined as*

$$\exists_x f \coloneqq f|_{x=1} + f|_{x=0}$$

for any $f \colon 2^X \to \mathbb{R}$.

Proposition 1 (Basic properties). *For any propositional formulas ϕ and ψ, and $a, b, c, d \in \mathbb{R}$,*

- $[\phi]^a_b = [\neg\phi]^b_a$;
- $c + [\phi]^a_b = [\phi]^{a+c}_{b+c}$;
- $c \cdot [\phi]^a_b = [\phi]^{ac}_{bc}$;
- $[\phi]^a_b \cdot [\phi]^c_d = [\phi]^{ac}_{bd}$;
- $[\phi]^1_0 \cdot [\psi]^1_0 = [\phi \wedge \psi]^1_0$.

And for any pair of pseudo-Boolean functions $f, g\colon 2^X \to \mathbb{R}$ and $x \in X$, $(fg)|_{x=i} = f|_{x=i} \cdot g|_{x=i}$ for $i = 0, 1$.

Definition 4 (old WMC instance). *An* old WMC instance *is a tuple (ϕ, X_I, X_P, w), where X_I is the set of indicator variables, X_P is the set of parameter variables (with $X_I \cap X_P = \emptyset$), ϕ is a propositional formula in CNF over $X_I \cup X_P$, and $w\colon X_I \cup X_P \cup \{\neg x \mid x \in X_I \cup X_P\} \to \mathbb{R}$ is the weight function. The* answer *of the instance is $\sum_{Y \models \phi} \prod_{Y \models l} w(l)$.*

Remark 1. Encodings such as `cd05` and `cd06` are *not* WMC encodings. Instead, they encode Bayesian network inference into instances of the *minimum-cardinality* WMC problem, where the answer is defined to be $\sum_{Y \models \phi,\ |Y|=k} \prod_{Y \models l} w(l)$, where $k = \min_{Y \models \phi,\ Y \neq \emptyset} |Y|$, if k exists, otherwise the answer is zero. This additional condition on model cardinality becomes necessary because these encodings eliminate clauses of the form $p \Rightarrow i$, where $p \in X_P$ is a parameter variable, and $i \in X_I$ is an indicator variable. Nonetheless, our transformation algorithm still works on such encodings, although the experimental results are discouraging because they use approximately twice as many indicator variables. For instance, each binary variable of a Bayesian network is encoded using two indicator variables while one would suffice.

Definition 5 (new WMC instance). *A* new WMC instance *is a tuple (F, X, ω), where X is the set of variables, F is a set of pseudo-Boolean functions $2^X \to \mathbb{R}$, and $\omega \in \mathbb{R}$ is the scaling factor. The* answer *of the instance is $\omega \cdot \left(\exists_X \prod_{f \in F} f \right) (\emptyset)$.*

References.

- Related work without publicly available implementations:
  - direct compilation to SDDs [1]
  - direct compilation to PSDDs, also eliminating parameter variables (a thesis)
  - maybe two more papers

Notes.

- Apparently, the DPMC paper already shows that taking the first offered decomposition tree is best.
- It is already well-known that WMC is FPT.
- Weights on literals other than the positive literals that correspond to variables in X_P are redundant as they either are equal to one or duplicate an already-defined weight.

- Mention that formulas, clauses, and models are all treated as sets.
- The constant is inspired by the `bklm16` encoding.
- We don't make a distinction between a real number and a (pseudo-Boolean) function that always returns that number.
- Throughout the paper, projection is assumed to have the lowest precedence (e.g., $\exists_x fg = \exists_x(fg)$ and not $(\exists_x f)g$).

2 Parameter Variable Elimination

Notes.

- Let X_P be the set of parameter variable and X_I be the set of indicator variables.
- Parameter variables are either taken from the LMAP file (for encodings produced by Ace) or assumed to be the variables that have both weights equal to 1.
- If a parameter variable in a clause is 'negated', we can ignore the clause. We assume that there are no clauses with more than one instance of parameter variables.
- The second **foreach** loop can be performed in constant time by representing ϕ' as a list and assuming that the two 'clauses' are adjacent in that list (and incorporating it into the first loop).
- The d map is constructed in $\mathcal{O}(|X_P| \log |X_P|)$ time (we want to use a data structure based on binary search trees rather than hashing).
- `rename` can be implemented in $\mathcal{O}(\log |X_P|)$ time.
- This may look like preprocessing, but all the transformations are local and thus can be incorporated into an encoding algorithm with no slowdown. In fact, if anything, the resulting algorithm would be slightly faster, as it would have less data to output.

2.1 Correctness

Theorem 1 (Early Projection [2, 3], verbatim). *Let X and Y be sets of variables. For all functions $f\colon 2^X \to \mathbb{R}$ and $g\colon 2^Y \to \mathbb{R}$, if $x \in X \setminus Y$, then $\exists_x(f \cdot g) = (\exists_x f) \cdot g$.*

Lemma 1. *For any pseudo-Boolean function $f\colon 2^X \to \mathbb{R}$, $(\exists_X f)(\emptyset) = \sum_{Y \subseteq X} f(Y)$.*

Proof. For base case, let $X = \{x\}$. Then

$$\exists_x f = f|_{x=1} + f|_{x=0},$$

by Definition 3 and so

$$(\exists_x f)(\emptyset) = f|_{x=1}(\emptyset) + f|_{x=0}(\emptyset)$$

Algorithm 1: WMC instance transformation

```
   Data: an (old-format) WMC instance $(\phi, X_I, X_P, w)$
   Result: a (new-format) WMC instance $(F, X, \omega)$
1  $F \leftarrow \emptyset$;
2  $\omega \leftarrow 1$;
3  let $d\colon X_P \to \mathbb{N}$ be defined as $p \mapsto |\{o \in X_P \mid o \leq p\}|$;
4  foreach clause $c \in \phi$ do
5      if $c \cap X_P = \{p\}$ for some $p$ and $w(p) \neq 1$ then
6          if $|c| = 1$ then
7              $\omega \leftarrow \omega \times w(p)$;
8          else
9              $F \leftarrow F \cup \left\{ \left[ \bigwedge_{l \in c \setminus \{p\}} \neg l \right]_1^{w(p)} \right\}$;
10     else if $\{p \mid \neg p \in c\} \cap X_P = \emptyset$ then
11         $F \leftarrow F \cup \{[c]_0^1\}$;
12 foreach indicator variable $v \in X_I$ do
13     if $\{[v]_1^p, [\neg v]_1^q\} \subseteq F$ for some $p$ and $q$ then
14         $F \leftarrow F \setminus \{[v]_1^p, [\neg v]_1^q\} \cup \{[v]_q^p\}$;
15 replace every variable $v$ in $F$ with rename($v$);
16 return $(F, X_I, \omega)$;
17 Function rename($v$):
18     $S \leftarrow \{u \in X_P \mid u \leq v\}$;
19     if $S = \emptyset$ then return $v$;
20     return $v - d(\max S)$;
```

by Definition 1. Then

$$f|_{x=1}(\emptyset) = f(\{x\}), \quad \text{and} \quad f|_{x=0}(\emptyset) = f(\emptyset)$$

by Definition 2. It follows that

$$(\exists_x f)(\emptyset) = f(\{x\}) + f(\emptyset) = \sum_{Y \subseteq \{x\}} f(Y)$$

as required. This easily extends to $|X| > 1$ by the definition of projection on sets of variables.

Proposition 2. *Let (ϕ, X_I, X_P, w) be an old WMC instance. Then*

$$\left(\left\{[c]^1_0 \mid c \in \phi\right\} \cup \left\{[x]^{w(x)}_{w(\neg x)} \,\middle|\, x \in X_I \cup X_P\right\}, X_I \cup X_P, 1\right) \tag{1}$$

is a new WMC instance with the same answer.

Proof. The answer of Eq. (2) is

$$\begin{aligned}\left(\exists_{X_I \cup X_P} \prod_{c \in \phi} [c]^1_0 \prod_{x \in X_I \cup X_P} [x]^{w(x)}_{w(\neg x)}\right)(\emptyset) &= \sum_{Y \subseteq X_I \cup X_P} \left(\prod_{c \in \phi} [c]^1_0 \prod_{x \in X_I \cup X_P} [x]^{w(x)}_{w(\neg x)}\right)(Y) \\ &= \sum_{Y \subseteq X_I \cup X_P} \underbrace{\left(\prod_{c \in \phi} [c]^1_0\right)(Y)}_{f} \underbrace{\left(\prod_{x \in X_I \cup X_P} [x]^{w(x)}_{w(\neg x)}\right)(Y)}_{g}\end{aligned}$$

by Definitions 1 and 5 and Lemma 1. Note that

$$f(Y) = \begin{cases} 1 & \text{if } Y \models \phi, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad g(Y) = \prod_{Y \models l} w(l),$$

which means that

$$\sum_{Y \subseteq X_I \cup X_P} f(Y) g(Y) = \sum_{Y \models \phi} \prod_{Y \models l} w(l)$$

as required.

Theorem 2 (Correctness). *Algorithm 1, when given an old WMC instance (ϕ, X_I, X_P, w), returns a new WMC instance with the same answer, provided the following conditions are satisfied:*

1. *for all indicator variables $i \in X_I$, $w(i) = w(\neg i) = 1$,*
2. *and either*
   (a) *for all parameter variables $p \in X_P$, there is a non-empty family of literals $(l_i)^n_{i=1}$ such that*

i. $w(\neg p) = 1$,
ii. $l_i \in X_I$ *or* $\neg l_i \in X_I$ *for all* $i = 1, \dots, n$,
iii. *and* $\{c \in \phi \mid p \in c$ *or* $\neg p \in c\} = \{p \vee \bigvee_{i=1}^{n} \neg l_i\} \cup \{l_i \vee \neg p \mid i = 1, \dots, n\}$;

(b) or for all parameter variables $p \in X_P$,
i. $w(p) + w(\neg p) = 1$,
ii. *for any clause* $c \in \phi$, $|c \cap X_P| \le 1$,
iii. *there is no clause* $c \in \phi$ *such that* $\neg p \in c$,
iv. *if* $\{p\} \in \phi$, *then there is no clause* $c \in \phi$ *such that* $c \neq \{p\}$ *and* $p \in c$,
v. *and for any* $c, d \in \phi$ *such that* $c \neq d$, $p \in c$ *and* $p \in d$, $\bigwedge_{l \in c \setminus \{p\}} \neg l \wedge \bigwedge_{l \in d \setminus \{p\}} \neg l$ *is false.*

Proof. By Proposition 2,

$$\left(\left\{ [c]^1_0 \mid c \in \phi \right\} \cup \left\{ [x]^{w(x)}_{w(\neg x)} \,\middle|\, x \in X_I \cup X_P \right\}, X_I \cup X_P, 1 \right) \tag{2}$$

is a new WMC instance with the same answer as the given old WMC instance. By Definition 5, its answer is

$$\left(\exists_{X_I \cup X_P} \prod_{c \in \phi} [c]^1_0 \prod_{x \in X_I \cup X_P} [x]^{w(x)}_{w(\neg x)} \right)(\emptyset) \tag{3}$$

Since both Conditions 2a and 2b ensure that each clause in ϕ has at most one parameter variable, we can partition ϕ into $\phi_* \coloneqq \{c \in \phi \mid \mathtt{Vars}(c) \cap X_P = \emptyset\}$ and $\phi_p \coloneqq \{c \in \phi \mid \mathtt{Vars}(c) \cap X_P = \{p\}\}$ for all $p \in X_P$. We can then use Theorem 1 to reorder Eq. (3) into

$$\left(\exists_{X_I} \left(\prod_{x \in X_I} [x]^{w(x)}_{w(\neg x)} \right) \left(\prod_{c \in \phi_*} [c]^1_0 \right) \prod_{p \in X_P} \exists_p [p]^{w(p)}_{w(\neg p)} \prod_{c \in \phi_p} [c]^1_0 \right)(\emptyset).$$

Let us first consider how the unfinished WMC instance (F, X_I, ω) after the loop on Lines 4 to 11 differs from Eq. (2). Note that Algorithm 1 leaves each $c \in \phi_*$ unchanged, i.e., adds $[c]^1_0$ to F. We can then fix an arbitrary $p \in X_P$ and let F_p be the set of functions added to F as a replacement of ϕ_p. It is sufficient to show that

$$\omega \prod_{f \in F_p} f = \exists_p [p]^{w(p)}_{w(\neg p)} \prod_{c \in \phi_p} [c]^1_0. \tag{4}$$

Note that under Condition 2a,

$$\bigwedge_{c \in \phi_p} c \equiv p \Leftrightarrow \bigwedge_{i=1}^{n} l_i$$

for some family of indicator variable literals $(l_i)_{i=1}^n$. Thus,

$$\exists_p [p]^{w(p)}_{w(\neg p)} \prod_{c \in \phi_p} [c]^1_0 = \exists_p [p]^{w(p)}_1 \left[p \Leftrightarrow \bigwedge_{i=1}^{n} l_i \right]^1_0.$$

If $w(p) = 1$, then

$$\exists_p [p]_1^{w(p)} \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 = \exists_p \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 = \left. \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 \right|_{p=1} + \left. \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 \right|_{p=0} \tag{5}$$

by Definition 3. Since for any input, $\bigwedge_{i=1}^n l_i$ is either true or false, exactly one of the two summands in Eq. (5) will be equal to one, and the other will be equal to zero, and so

$$\left. \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 \right|_{p=1} + \left. \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 \right|_{p=0} = 1,$$

where 1 is a pseudo-Boolean function that always returns one. On the other side of Eq. (4), since $F_p = \emptyset$, and ω is unchanged, we get $\omega \prod_{f \in F_p} f = 1$, and so Eq. (4) is satisfied under Condition 2a when $w(p) = 1$.

If $w(p) \neq 1$, then

$$F_p = \left\{ \left[\bigwedge_{i=1}^n l_i \right]_1^{w(p)} \right\},$$

and $\omega = 1$, and so we want to show that

$$\left[\bigwedge_{i=1}^n l_i \right]_1^{w(p)} = \exists_p [p]_1^{w(p)} \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 .$$

By Definition 3,

$$\begin{aligned}
\exists_p [p]_1^{w(p)} \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 &= \left. \left([p]_1^{w(p)} \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 \right) \right|_{p=1} + \left. \left([p]_1^{w(p)} \left[p \Leftrightarrow \bigwedge_{i=1}^n l_i \right]_0^1 \right) \right|_{p=0} \\
&= w(p) \cdot \left[\bigwedge_{i=1}^n l_i \right]_0^1 + \left[\bigwedge_{i=1}^n l_i \right]_1^0 = \left[\bigwedge_{i=1}^n l_i \right]_1^{w(p)}
\end{aligned}$$

as required. This finishes the proof of the correctness of the first 'foreach' loop under Condition 2a.

Now let us assume Condition 2b. We still want to prove Eq. (4). If $w(p) = 1$, then $F_p = \emptyset$, and $\omega = 1$, and so the left-hand side of Eq. (4) is equal to one. Then the right-hand side is

$$\exists_p [p]_0^1 \prod_{c \in \phi_p} [c]_0^1 = \exists_p \left[p \wedge \bigwedge_{c \in \phi_p} c \right]_0^1 = \exists_p [p]_0^1 = 0 + 1 = 1$$

by Proposition 1 and Definition 3, and because $p \in c$ for every clause $c \in \phi_p$.

If $w(p) \neq 1$, and $\{p\} \in \phi_p$, then, by Condition 2(b)iv, $\phi_p = \{\{p\}\}$, and Algorithm 1 produces $F_p = \emptyset$ and $\omega = w(p)$, and so

$$\omega \prod_{f \in F_p} f = w(p).$$

Whereas on the other side of Eq. (4),

$$\exists_p [p]^{w(p)}_{w(\neg p)} [p]^1_0 = \exists_p [p]^{w(p)}_0 = w(p)$$

by Proposition 1 and Definition 3.

The only remaining case is when $w(p) \neq 1$ and $\{p\} \notin \phi_p$. Then $\omega = 1$, and

$$F_p = \left\{ \left[\bigwedge_{l \in c \setminus \{p\}} \neg l \right]^{w(p)}_1 \middle| \, c \in \phi_p \right\},$$

so need to show that

$$\prod_{c \in \phi_p} \left[\bigwedge_{l \in c \setminus \{p\}} \neg l \right]^{w(p)}_1 = \exists_p [p]^{w(p)}_{1-w(p)} \prod_{c \in \phi_p} [c]^1_0.$$

We can rearrange the right-hand-side as

$$\begin{aligned}
\exists_p [p]^{w(p)}_{1-w(p)} \prod_{c \in \phi_p} [c]^1_0 &= \exists_p [p]^{w(p)}_{1-w(p)} \left[\bigwedge_{c \in \phi_p} c \right]^1_0 = \exists_p [p]^{w(p)}_{1-w(p)} \left[p \vee \bigwedge_{c \in \phi_p} c \setminus \{p\} \right]^1_0 \\
&= w(p) + (1 - w(p)) \left[\bigwedge_{c \in \phi_p} c \setminus \{p\} \right]^1_0 = w(p) + \left[\bigwedge_{c \in \phi_p} c \setminus \{p\} \right]^{1-w(p)}_0 \\
&= \left[\bigwedge_{c \in \phi_p} c \setminus \{p\} \right]^1_{w(p)} = \left[\neg \bigwedge_{c \in \phi_p} c \setminus \{p\} \right]^{w(p)}_1 = \left[\bigvee_{c \in \phi_p} \neg (c \setminus \{p\}) \right]^{w(p)}_1 \\
&= \left[\bigvee_{c \in \phi_p} \neg \bigvee_{l \in c \setminus \{p\}} l \right]^{w(p)}_1 = \left[\bigvee_{c \in \phi_p} \bigwedge_{l \in c \setminus \{p\}} \neg l \right]^{w(p)}_1
\end{aligned}$$

by Proposition 1 and Definition 3. By Condition 2(b)v, $\bigwedge_{l \in c \setminus \{p\}} \neg l$ can be true for at most one $c \in \phi_p$, and so

$$\left[\bigvee_{c \in \phi_p} \bigwedge_{l \in c \setminus \{p\}} \neg l \right]^{w(p)}_1 = \prod_{c \in \phi_p} \left[\bigwedge_{l \in c \setminus \{p\}} \neg l \right]^{w(p)}_1$$

which is exactly what we needed to show. This ends the proof that the first loop of Algorithm 1 preserves the answer under both Condition 2a and Condition 2b. Finally, the loop on Lines 12 to 14 of Algorithm 1 replaces $[v]_1^p[\neg v]_1^q$ with $[v]_q^p$ (for some $v \in X_I$ and $p, q \in \mathbb{R}$), but, of course,

$$[v]_1^p[\neg v]_1^q = [v]_1^p[v]_q^1 = [v]_q^p$$

by Proposition 1.

Notes.

- This just says that p is equivalent to a conjunction.
- The first group of conditions applies to `d02`, while the second group applies to `bklm16`.
- For `cd05` and `cd06`, condition $2bi$ should be replaced with $w(\neg p) = 1$.
- Benefits of having this proof in the paper:
  - It puts all encodings on a common ground.
  - It illustrates the convenience of our notation for reasoning about (certain types of) pseudo-Boolean functions.
  - It's too big and too important to be left for the appendix.
- Processors:
  - Intel(R) Xeon(R) Gold 6138 CPU @ 2.00GHz
  - Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz
  - Intel(R) Xeon(R) CPU E7-4820 v2 @ 2.00GHz
- Table captions should be placed above the tables.
- Figures should preferably be in EPS format.
- the environments 'theorem', 'definition', 'lemma', 'proposition', 'corollary', 'remark', and 'example' are defined in the LLNCS documentclass as well.

TODO

- Come up with better names for old/new WMC instances.
- Do I need to formally consider extending a pseudo-Boolean function to a bigger domain?
- Define scalar operations on ADDs.
- Define projection of a set of variables.
- add parentheses around products
- maybe give a name to (2) since it's wrong to call it an equation
- check if each condition is actually used. Maybe turn this into a paragraph that gives an overview of the proof.
- condition 1 is necessary in both cases because we're ignore the weights of indicator variables
- maybe do more summations over domains in proofs

References

1. Choi, A., Kisa, D., Darwiche, A.: Compiling probabilistic graphical models using sentential decision diagrams. In: van der Gaag, L.C. (ed.) Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 12th European Conference, ECSQARU 2013, Utrecht, The Netherlands, July 8-10, 2013. Proceedings. Lecture Notes in Computer Science, vol. 7958, pp. 121–132. Springer (2013). https://doi.org/10.1007/978-3-642-39091-3_11
2. Dudek, J.M., Phan, V., Vardi, M.Y.: ADDMC: weighted model counting with algebraic decision diagrams. In: The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020. pp. 1468–1476. AAAI Press (2020), https://aaai.org/ojs/index.php/AAAI/article/view/5505
3. Dudek, J.M., Phan, V.H.N., Vardi, M.Y.: DPMC: weighted model counting by dynamic programming on project-join trees. In: Simonis, H. (ed.) Principles and Practice of Constraint Programming - 26th International Conference, CP 2020, Louvain-la-Neuve, Belgium, September 7-11, 2020, Proceedings. Lecture Notes in Computer Science, vol. 12333, pp. 211–230. Springer (2020). https://doi.org/10.1007/978-3-030-58475-7_13